

HISTORY – INSTITUTE FOR EARLY MODERN HISTORY/ GHENTCDH RESEARCH LAB

Christel Annemieke Romein

ENTANGLED HISTORIES – MAKING ORDINANCES SEARCHABLE (±1500-1800)

Aims of the Project:

This project

- (1) improves the currently applied Optical Character Recognition (OCR)-technique to a much higher recognition-standard with Handwritten Text Recognition (HTR).
- (2) enhances readability by systematically segmenting individual texts, recognising text-sections – beginning or end, columns, titles, dates, summaries, the body of the text.
- (3) applies a standard categorisation (metadata) with a machine-learned algorithm.

Underlying hypothesis

When problems arose, small 'states' had to act swiftly. Hence, I assume that they may have adopted – parts of – successful legislation from neighbouring 'states'. Hence 'entangled histories'.

Improved searchability

Improving the searchability allows to search for e.g. keywords, dates and titles. These can be exported and used to search the books of ordinances more quickly and visualise output.

What metadata (categories) is added?

The Max-Planck-Institute für europäische Rechtsgeschichte created a hierarchical structure of categories within their Policygesetzgebung's project (Karl Härter, Michael Stolleis). These categories are going to be applied as much as possible to enable international comparisons.

Relevance and implications

- Improved accessibility to a huge amount of normative texts from the early modern era, allowing longitudinal research to 'common' problems.
- A frequently used resource, now being placed in context (e.g. able to find similar texts in other areas and periods) and accessible through metadata.

Sources and Readability

This project uses the books of ordinances ('plakkaatboeken') that were issued by the various governments within the Low Countries (North and South).

These have been digitised within the Google Books project but their OCR-quality is poor (est. 40% Character Error Rate (CER)).

This project aims to improve the OCR-quality to a <5% CER in order to make them better searchable for researchers and computers.

The readability of the texts is improved by using Handwriting Text Recognition (HTR) tool **Transkribus**.



Groot Plakkaatboeken – Museum De Roos (Geertruidenberg)
<http://www.museumderoos.nl/index.php?menuitemID=106&taalID=2>

Image: Habsburg Netherlands; David Descamps
https://commons.wikimedia.org/wiki/File:Spanish_Netherlands.svg#/media/File:Spainish_Netherlands.svg